



## Genetic analysis on infrared-predicted milk minerals for Danish dairy cattle

R. M. Zaalberg,<sup>1\*</sup> N. A. Poulsen,<sup>2</sup> H. Bovenhuis,<sup>3</sup> J. Sehested,<sup>4</sup> L. B. Larsen,<sup>2</sup> and A. J. Buitenhuis<sup>1</sup>

<sup>1</sup>Center for Quantitative Genetics and Genomics, Aarhus University, Blichers Allé 20, DK-8830 Tjele, Denmark

<sup>2</sup>Department of Food Science, Aarhus University, Agro Food Park 48, 8200 Aarhus N, Denmark

<sup>3</sup>Animal Breeding and Genomics, Wageningen University and Research, 6700AH, Wageningen, The Netherlands

<sup>4</sup>Department of Animal Science, Aarhus University, Blichers Allé 20, DK-8830 Tjele, Denmark

### ABSTRACT

A group of milk components that has shown potential to be predicted with milk spectra is milk minerals. Milk minerals are important for human health and cow health. Having an inexpensive and fast way to measure milk mineral concentrations would open doors for research, herd management, and selective breeding. The first aim of this study was to predict milk minerals with infrared milk spectra. Additionally, milk minerals were predicted with infrared-predicted fat, protein, and lactose content. The second aim was to perform a genetic analysis on infrared-predicted milk minerals, to identify QTL, and estimate variance components. For training and validating a multibreed prediction model for individual milk minerals, 264 Danish Jersey cows and 254 Danish Holstein cows were used. Partial least square regression prediction models were built for Ca, Cu, Fe, K, Mg, Mn, Na, P, Se, and Zn based on 80% of the cows, selected randomly. Prediction models were externally validated with 8 herds based on the remaining 20% of the cows. The prediction models were applied on a population of approximately 1,400 Danish Holstein cows with 5,600 infrared spectral records and 1,700 Danish Jersey cows with 7,200 infrared spectral records. Cows from this population had 50k imputed genotypes. Prediction accuracy was good for P and Ca, with external  $R^2 \geq 0.80$  and a relative prediction error of 5.4% for P and 6.3% for Ca. Prediction was moderately good for Na with an external  $R^2$  of 0.63, and a relative error of 18.8%. Prediction accuracies of milk minerals based on infrared-predicted fat, protein, and lactose content were considerably lower than those based on the infrared milk spectra. This shows that the milk infrared spectrum contains valuable information on milk minerals, which is currently not used. Heritability for infrared-predicted Ca, Na, and P varied from low (0.13) to moderate (0.36). Several QTL for

infrared-predicted milk minerals were observed that have been associated with gold standard milk minerals previously. In conclusion, this study has shown infrared milk spectra were good at predicting Ca, Na, and P in milk. Infrared-predicted Ca, Na, and P had low to moderate heritability estimates.

**Key words:** mid infrared, spectroscopy, novel phenotype

### INTRODUCTION

Development of prediction models with the use of infrared milk spectra is a hot topic in dairy science (e.g., Soyeurt et al., 2006, 2011; De Marchi et al., 2018). Infrared milk spectra are promising, because spectroscopic analysis of milk is inexpensive and fast, and when milk spectra are stored, they can be applied for prediction of new phenotypes at any time (De Marchi et al., 2014). Currently, infrared milk spectra are used for routine analysis of milk components, such as fat, protein, and lactose. These milk components can be predicted from infrared milk spectra with high precision, having prediction accuracies up to 1 (Luinge et al., 1993; Soyeurt et al., 2006). Infrared milk spectra have also been investigated for prediction of phenotypes such as concentration of milk fatty acids (Soyeurt et al., 2006; Eskildsen et al., 2014), concentration of BHB and acetone (Heuer et al., 2001; Van der Drift et al., 2012), energy status (McParland et al., 2011), methane emission (Dehareng et al., 2012), minerals (Soyeurt et al., 2009), milk metabolites (Zaalberg et al., 2020), and even genotypes (Rutten et al., 2011).

Infrared (predicted) phenotypes could be used for herd management or for selective breeding. Before infrared phenotypes are used for breeding purposes, several points should be discussed. First, they should be of direct or indirect economic interest for either farmer or breeder. Second, accurate infrared phenotypes are desirable when using infrared phenotypes on individual cow level, such as using infrared phenotypes as a tool for deciding on the use of veterinary treatment. However, breeding values of bulls are estimated based on

Received September 14, 2020.

Accepted March 26, 2021.

\*Corresponding author: [roos.zaalberg@qgg.au.dk](mailto:roos.zaalberg@qgg.au.dk)

many (repeated) daughter records, which allows for correcting noise in phenotypes (Bovenhuis et al., 2018). Infrared phenotypes that can be predicted with a lower accuracy can, therefore, still be valuable in breeding programs (Bovenhuis et al., 2018). Third, to improve the future dairy cattle population, the infrared phenotype should meet 3 conditions: (1) there has to be phenotypic variation for the infrared phenotype, (2) the infrared phenotype should be heritable, and (3) it should be genetically correlated with the gold standard phenotype.

A group of milk components that potentially meets these 3 requirements is milk minerals. Milk is an important source for minerals in human nutrition, including Ca, P, Mg, Se, An, and K (Haug et al., 2007; Gaucheron, 2013). Furthermore, elevated levels of sodium in milk can be an indicator for mastitis (Gaucheron, 2013). Tracing mastitis at an early stage is important, both from an animal welfare and financial point of view. Milk minerals are also relevant for technological aspects of milk processing (Lucey and Fox, 1993; Jensen et al., 2012). Calcium phosphate influences physicochemical stability of casein in milk, cheese-making properties, and milk quality during storage (Lucey and Fox, 1993; Gaucheron, 2005; Haug et al., 2007). Changing levels of calcium influenced stability of milk to heat treatment (Deeth and Lewis, 2015). This link between milk minerals and milk properties is supported by the genetic correlations that were observed between milk minerals and cheese-making properties (Sanchez et al., 2018, 2019).

Measuring mineral concentration in milk is a costly and time-consuming process. Infrared prediction of milk minerals could, therefore, be a valuable and inexpensive addition to the collection of milk traits that are currently available for farm management and breeding schemes. For several milk minerals, accurate infrared prediction models have been developed (e.g., Soyeurt et al., 2009; Visentin et al., 2016).

Genetic analysis of infrared-predicted phosphorus and calcium revealed several QTL in Montbéliarde cows (Sanchez et al., 2019). Additionally, QTL have been presented for 10 gold standard minerals in Danish Holstein and Danish Jersey (Buitenhuis et al., 2015). Heritability estimates for infrared-predicted minerals vary across breeds and countries and are overall low to moderately high (e.g., Toffanin et al., 2015; Visentin et al., 2016; Sanchez et al., 2018). To effectively apply infrared-predicted minerals in Danish breeding schemes, we will have to better understand the genetic background of these phenotypes in Danish dairy cattle.

The first aim of this study was to predict minerals in milk using infrared milk spectra for 2 Danish dairy cattle breeds. Additionally, milk minerals were predicted based on major milk components predicted

with infrared milk spectra (fat, protein, and lactose). Second, a genetic analysis was performed on predicted milk minerals to identify QTL and estimate variance components.

## MATERIALS AND METHODS

### *Building the Infrared Prediction Model for Milk Minerals*

**Population for Building the Infrared Prediction Model.** To build the infrared prediction model, a multibreed population was used that consisted of 254 Danish Holstein cows and 281 Danish Jersey cows. The 254 Danish Holstein cows were between 129 and 229 DIM, and originated from 19 herds. Milk samples of at least 0.5 L were collected from May until June 2013. The 281 Danish Jersey cows were between 130 and 252 DIM, and originated from 20 herds. Milk samples of at least 0.5 L were collected from February until April 2010. All cows used in this study were in parity 1, 2, or 3. All samples were morning milk samples. Milk samples were placed on ice after collection, and transported to the laboratory for analysis (Poulsen et al., 2012). At the laboratory, one fresh milk sample was taken and used for spectral analysis, and one sample was frozen for analysis of minerals.

**Infrared Milk Spectra.** For both 254 Danish Holstein cows and 281 Danish Jersey cows infrared spectral analysis on fresh milk samples was performed using the MilkoScan FT2 (FOSS Analytical A/S) at the Aarhus University research facility (Tjele, Denmark). Transmittance values were obtained for 1,060 individual wavenumbers of the infrared spectral region 5,008 through 925  $\text{cm}^{-1}$ . Transmittance values of individual wavenumbers were transformed to absorbance values using:

$$\text{Absorbance} = \text{Log}_{10} \left( \frac{1}{\text{Transmittance}} \right). \quad [1]$$

Two infrared spectral regions that interact with water molecules were removed: 5,008 through 3,008  $\text{cm}^{-1}$ , and wavenumber 1,700 through 1,600  $\text{cm}^{-1}$ . The remaining 534 wavenumbers were used to build the mid-infrared (MIR) prediction model.

To remove spectral outliers, standardized Mahalanobis distances (MD) were calculated. A principal component analysis (PCA) was performed on the infrared spectra of 535 cows (254 Danish Holstein and 281 Danish Jersey cows). The PCA was done in R (version 3.4.1; <http://www.r-project.org>), using the “prcomp” and “Mahalanobis” function from the “stats” package.

Each milk record had absorbance values for 534 wave-numbers. The standardized MD was calculated as the average MD of the score values for the first 3 principle components. Records with a standardized MD of >5 were considered as outliers. Based on this criterion, 1 Danish Holstein and 8 Danish Jersey records were removed.

**Milk Components.** Fat, protein, and lactose content were analyzed in fresh, full-fat milk by infrared spectroscopy (Milkoscan FT2, FOSS Analytical A/S). The instrument was routinely calibrated according to these milk parameters by use of calibration samples provided by Eurofins Scientific.

**Gold Standard Minerals.** In 2014, 10 minerals (Ca, Cu, Fe, K, Mg, Mn, Na, P, Se, and Zn) were extracted from the unfrozen skim milk sample by acid sonication and identified using inductively coupled plasma MS. For identifying minerals, inductively coupled plasma MS (ICP-MS) was used (Cava-Montesinos et al., 2005). For an elaborate description of the method, see Buitenhuis et al. (2015). No milk samples had mineral concentrations that deviated more than 3 standard deviations from the mean.

**Training and Validating the Infrared Prediction Models.** Partial least squares regression (PLSR) was used for predicting gold standard minerals with infrared milk spectra. Infrared prediction models for milk minerals were built using the package “pls” in R (version 3.4.1; <http://www.r-project.org>).

For each mineral separately, a multibreed prediction model was built based on infrared milk spectra from 253 Danish Holstein cows and 273 Danish Jersey cows. The procedure to build the MIR-based prediction model consisted of 3 steps. During the first step, the

aim was to determine the optimal number of latent variables that should be included in the final prediction model. To train the prediction model, 80% of the Danish Holstein cows and 80% of the Danish Jersey cows were randomly selected. The remaining 20% for both breeds were used to validate the prediction model. This process was repeated 100 times. These 100 repeats were used to calculate the average coefficient of determination ( $R^2$ ), and the average residual means squared error of prediction (RMSEP). The average RMSEP and  $R^2$  were analyzed to determine the optimal number of latent variables. Once both the RMSEP and  $R^2$  improved by  $\leq 5\%$ , the optimum number of latent variables was selected (Zaalberg et al., 2020). During the second step, we used the optimum number of latent variables from step 1, and performed an external validation to test the performance of the prediction model. The validation population consisted of 4 herds with 51 Danish Holstein cows (20%), 4 herds with 53 Danish Jersey cows (20%). To select herds for validating the model, a prediction model was built with the remainder animals, and applied on the validation population. The prediction accuracy ( $R^2$ ) and RMSEP were calculated. During the third and final step, the optimal number of latent variables from step 1 was used to build a final multibreed prediction model that was trained with all 254 Danish Holstein and 265 Danish Jersey (Table 1). This model was used to predict milk mineral concentration for a larger genotyped study population (see next section of Materials and Methods).

**Prediction Models Based on Infrared-Predicted Fat, Protein, and Lactose.** The same procedure as described for infrared milk spectra was used to build prediction models based on the following infrared-

**Table 1.** Accuracy ( $R^2$ ), the residual means squared error of prediction (RMSEP), and the percentage of RMSEP relative to the gold standard mean (Error%) for prediction models<sup>1</sup> based on infrared milk spectra

Mineral	Training population, <sup>2</sup> n		Gold standard concentration		Infrared prediction model			
	DH	DJ	Mean (SD)		LV <sup>3</sup>	$R^2$	RMSEP	Error%
			DH	DJ				
Calcium, g/L	253	273	1.24 (0.13)	1.46 (0.15)	10	0.80	0.08	6.3
Copper, g/L	253	273	28.50 (12.80)	51.50 (23.30)	3	0.39	18.00	41.7
Iron, mg/L	253	273	0.16 (0.03)	0.18 (0.05)	3	0.25	0.04	22.0
Potassium, g/L	253	273	1.47 (0.11)	1.33 (0.11)	17	0.33	0.11	7.5
Magnesium, g/L	253	273	0.11 (0.01)	0.12 (0.01)	5	0.46	0.01	8.1
Manganese, $\mu\text{g/L}$	253	273	16.10 (5.00)	30.60 (9.10)	5	0.20	8.00	30.8
Sodium, g/L	253	273	0.35 (0.07)	0.41 (0.13)	9	0.63	0.07	18.8
Phosphorus, g/L	253	273	0.72 (0.08)	0.88 (0.08)	12	0.87	0.04	5.4
Selenium, $\mu\text{g/L}$	253	273	6.80 (1.30)	10.60 (2.00)	5	0.32	2.00	20.8
Zinc, mg/L	253	273	3.38 (0.56)	4.67 (0.80)	2	0.51	0.67	16.1

<sup>1</sup>A multibreed training population of Danish Holstein and Danish Jersey was used to build the prediction model. The prediction model was externally validated with 4 Danish Holstein herds and 4 Danish Jersey herds.

<sup>2</sup>All animals have one milk record. DH = Danish Holstein, initial population size 254; DJ = Danish Jersey, initial population size 281.

<sup>3</sup>LV = latent variables.

predicted milk components: fat, protein, lactose, fat + protein, and fat + protein + lactose. For models based on infrared-predicted milk components, the number of latent variables was the same as the number of milk components that was used to train the prediction model (1, 2, or 3 components).

### Genetic Parameter Estimation for Predicted Minerals

Genetic analyses were performed on milk mineral concentration predicted with infrared milk spectra, and milk mineral concentration predicted with infrared-predicted fat + protein. The genetic analysis was only performed on those minerals, which could be predicted using infrared milk spectra with an  $R^2 > 0.60$  (see step 2 in the Training and Validating the Infrared Prediction Models section).

**Genotyped Study Populations.** Two genotyped populations were used in this study, which will be referred to as study populations. The first study population was a Danish Holstein population that consisted of 3,211 cows from 308 herds. The second was a Danish Jersey study population that consisted of 3,360 cows from 153 herds. Morning milk samples were collected during routine milk recordings in the period of October 2015 through September 2016. When more than one milk sample was available for a given day for one cow, all milk samples of that cow for the day in question were excluded. To match the lactation stage of the cows that were used to build the prediction models, only milk records that were collected between 100 and 250 DIM were included in the analysis. Herds which had milk records from less than 5 cows were excluded. After data filtering, 1,436 Danish Holstein from 88 herds with 5,727 milk records, and 1,751 Danish Jersey from 52 herds with 7,307 milk records were left for analysis. Cows had between 1 and 10 milk records, with an average of 4 records.

**Infrared Milk Spectra.** All milk samples from the genotyped study populations were analyzed with the same MilkoScan FT6000 spectrometer (FOSS Analytical A/S) at the Eurofins-Steins laboratory (Vejen, Denmark). The spectrometer was calibrated every 2 wk. Transmittance values were obtained for 1,060 individual wavenumbers of the infrared spectral region 5,008 through 925  $\text{cm}^{-1}$ . These transmittance values were treated as described previously. Spectral outliers were removed by studying PLSR Hotelling  $T^2$  versus  $Q$ -residual plots, and leverage versus studentized residual plots (Zaalberg et al., 2019).

**Milk Components.** The content of fat, and protein in milk samples was predicted from infrared milk spectra during the standard analysis of milk samples by

Eurofins-Steins laboratory using prediction equations provided by FOSS Analytical A/S.

**Genotypes.** Cows from the study populations were genotyped with the EuroG10K custom SNP chip (10K-chip). The 10K-chip is a selection of SNPs that was selected from the whole-genome-sequencing data of bulls from the 1,000-bull-genome project. Selection of SNPs was based on association to production traits. The 10K-genotypes were imputed to the 50K using BEAGLE 4 (Browning and Browning, 2016). A reference population that was genotyped for Illumina 50K BovineSNP50 v.2 BeadChip (50K-chip; Illumina Inc.) was used. For Danish Holstein, the reference population consisted of 4,000 cows, and for Danish Jersey, the reference population consisted of 4,576 cows. Reference cows were genotyped with the Illumina 50K BovineSNP50 v.2 BeadChip (50K-chip; Illumina Inc.). For the genetic analysis, only SNPs that were located on autosomes were selected, SNPs that had more than 40% missing genotypes were excluded, and SNPs with a MAF  $< 1\%$  were removed. After these restrictions, for Danish Holstein 43,807 SNPs, and for Danish Jersey 39,235 SNPs were available for the genetic analysis.

**Infrared-Predicted Minerals for Study Populations.** Minerals were predicted with infrared milk spectra using the final prediction models as described in the section Training and Validating the Infrared Prediction Model. Milk records with predicted mineral concentrations that deviated by more than 3 standard deviations from the mean predicted mineral concentration were removed. The total number of cows after removal of outliers are presented in Appendix Table A1.

**Model Description for Genetic Analysis Study Populations.** Infrared-predicted minerals were analyzed for the 2 study populations separately. The genetic analysis was done using a Bayesian approach with the use of the software Bayz (<http://www.bayz.biz/>). Infrared-predicted minerals were analyzed with the following model:

$$y_{ijk} = \mu + Parity_i + \beta_1 DIM_{ijk} + \beta_2 e^{-0.05 DIM_{ijk}} + HYS_j + CowA_k + CowPE_k + e_{ijk}, \quad [2]$$

where  $y_{ijk}$  is the infrared mineral concentration for one milk record;  $\mu$  is the mean;  $Parity_i$  is a fixed effect ( $i = 1$  or  $2$ );  $\beta_1 DIM_{ijk}$  and  $\beta_2 e^{-0.05 DIM_{ijk}}$  adjust for the lactation stage (Wilmink function; Wilmink, 1987), where  $DIM_{ijk}$  is the DIM expressed in years for one milk record. For  $Parity$  and  $DIM$ , we assumed a uniform prior distribution, where values ranged from  $-\infty$  to  $+\infty$ ;  $HYS_j$  is a herd-year-season effect (autumn from Octo-

ber 1 to December 31, 2015, winter from January 1 to March 31, 2016, spring from April 1 to June 30, 2016, and summer from July 1 to September 30, 2016), for which a normal prior distribution was assumed, where  $HYS \sim N(0, \sigma_{HYS}^2)$ ;  $CowPE_k$  is the permanent environmental effect for cow  $k$ , for which a normal prior distribution was assumed, where  $CowPE \sim N(0, \sigma_{PE}^2)$ ; and  $e_{ijk}$  is the residual variance, for which a normal prior distribution was assumed, where  $e \sim N(0, \sigma_e^2)$ .  $CowA_k$  was modeled using SNP data with the following model:

$$CowA_k = \sum_m a_m g_{km}, \quad [3]$$

where  $CowA_k$  is the additive genetic value for cow  $k$ , modeled as the sum of additive SNP effects  $a_m$  of SNP  $m$  times centered allele dosages  $g_{km}$  (0, 1, or 2) for SNP  $m$  of cow  $k$ . A BayesA model was used for the additive SNP effects with  $a_m \sim N(0, \sigma_{am}^2)$ . All  $\sigma_{am}^2$  have a common scaled inverse-chi-squared distribution, and a scale parameter estimated from the data with a uniform prior distribution. The additive genetic variance modeled by all SNPs was  $\sigma_A^2 = var(CowA)$ . This parameter was computed by evaluating  $var(CowA)$  over the samples of the Monte Carlo Markov Chain. All SNP variances were estimated simultaneously.

**Variance Components and Heritability.** Heritability was estimated for each breed and mineral separately as:

$$h_{SNP}^2 = \frac{\sigma_A^2}{\sigma_{HYS}^2 + \sigma_A^2 + \sigma_{PE}^2 + \sigma_e^2}. \quad [4]$$

The variation explained by herd-year-season was calculated as:

$$HYS^2 = \frac{\sigma_{HYS}^2}{\sigma_{HYS}^2 + \sigma_A^2 + \sigma_{PE}^2 + \sigma_e^2}, \quad [5]$$

where  $\sigma_A^2$  is the additive genetic variance;  $\sigma_{HYS}^2$  is the herd-year-season variance;  $\sigma_{PE}^2$  is the permanent environmental variance; and  $\sigma_e^2$  is the residual variance.

**Identifying Quantitative Trait Loci.** Based on estimates for  $\sigma_A^2$  obtained with Model 3, genomic EBV (GEBV) were calculated for groups of 100 consecutive SNPs (Gebreyesus et al., 2017). For creating groups of SNPs, 5 starting points were used on each chromosome: SNP 1, SNP 21, SNP 41, SNP 61, and SNP 81. Conse-

quently, groups of SNPs have a maximum overlap of 80 SNPs, but never more. All SNPs within an SNP group were located on the same chromosome. Genomic EBV were estimated using Bayz and gBayz software (<http://www.bayz.biz/>).

For each SNP group the percentage of explained additive genetic variance was calculated as:

$$\% \sigma_{A,i}^2 = \frac{\sigma_{SNP,i}^2}{\sigma_A^2} \times 100\%, \quad [6]$$

where  $\sigma_{SNP,i}^2$  is the variance of the milk mineral explained by SNP group  $i$ , and  $\sigma_A^2$  is the additive genetic variance of the milk mineral.

A significant QTL was detected when a SNP group explained >0.5% of the additive genetic variation of the milk mineral. The lead SNP group is the SNP group that explained the highest proportion of genetic variation within the genomic region of the QTL. Gene annotation was performed in Ensemble (92), with the UMD3.1 assembly, and the Variant Effect Predictor function (McLaren et al., 2016).

## RESULTS AND DISCUSSION

This study aimed at predicting milk minerals from Fourier transform infrared milk spectra, and to study genetic parameters and quantitative trait loci for infrared-predicted milk minerals. In addition, milk mineral concentration was predicted based on already available infrared-predicted milk components (fat, protein, and lactose).

### Prediction Models

**Gold Standard Milk Minerals.** Gold standard milk mineral concentrations for the training populations are presented in Table 1. The milk mineral concentrations observed in the current study were very similar to those from previous studies, even though gold standard methods varied across studies (e.g., Soyeurt et al., 2009; Bonfatti et al., 2016; Visentin et al., 2016). The process of freezing can change milk mineral concentration due to interactions between milk minerals and milk fat (Kent et al., 2009). Milk samples in the current study were skimmed before freezing. Therefore, we expect no effect of freezing on milk mineral concentrations.

**Calcium.** Infrared milk spectra accurately predicted milk Ca concentration (Table 1). Infrared prediction model performance was similar to results reported in other studies, yet the RMSEP of the current study

**Table 2.** Accuracy ( $R^2$ ), the residual means squared error of prediction (RMSEP), and the percentage of RMSEP relative to the gold standard mean (Error%) for milk minerals predicted with infrared-predicted fat, protein, lactose, or a combination of milk components<sup>1</sup> (Table 1)

Infrared-predicted milk component	LV <sup>2</sup>	Calcium			Sodium			Phosphorus		
		$R^2$	RMSEP	Error%	$R^2$	RMSEP	Error%	$R^2$	RMSEP	Error%
Fat	1	0.42	0.14	10.4	0.02	0.11	29.4	0.44	0.09	10.7
Protein	1	0.58	0.12	8.9	0.03	0.11	29.2	0.58	0.07	9.3
Lactose	1	0.07	0.18	13.2	0.48	0.83	22.7	0.03	0.11	14.2
Fat + protein	2	0.59	0.12	8.7	0.03	0.11	29.2	0.60	0.07	9.1
Fat + protein + lactose	3	0.60	0.12	8.6	0.49	0.08	21.7	0.65	0.07	8.5

<sup>1</sup>Results are only presented for minerals that had an infrared prediction model with an accuracy >0.60.

<sup>2</sup>LV = latent variables.

was lower (Soyeurt et al., 2009; Toffanin et al., 2015; Bonfatti et al., 2016; Visentin et al., 2016; Franzoi et al., 2019). Infrared-predicted fat and protein predicting Ca in milk (Table 2) with moderate accuracy (0.59). Infrared-predicted lactose was a poor predictor for Ca content in milk. This is in line with the phenotypic and genetic correlations presented by other studies (Toffanin et al., 2015; Costa et al., 2019; Visentin et al., 2019).

**Sodium.** Infrared milk spectra were moderate at predicting milk Na concentration (Table 1). The observed  $R^2$  was similar to those reported in other studies (Soyeurt et al., 2009; Visentin et al., 2016; Franzoi et al., 2019; Sanchez et al., 2019). Furthermore, infrared-predicted lactose content had a strong predictive ability with regards to milk Na (Table 2). This is as expected, because lactose and Na in milk are inversely related, which is necessary to maintain the osmotic equilibrium in milk (Keenan and Patton, 1995). This is in line with the strong negative phenotypic and genetic correlations between lactose content and Na in milk (Costa et al., 2019).

**Phosphorus.** Out of all minerals, infrared milk spectra were best at predicting milk P concentrations. The external prediction accuracy of 0.87 (Table 1) was high compared with results presented by other studies (Soyeurt et al., 2009; Toffanin et al., 2015; Visentin et al., 2016; Franzoi et al., 2019; Sanchez et al., 2019). Both infrared-predicted protein and fat were moderate at predicting P in milk. These results are in line with previously described predictive abilities of milk components for P in milk (Bovenhuis et al., 2018). Furthermore, strong phenotypic and genetic correlation between P and protein content in mid-lactation milk has been presented before (Toffanin et al., 2015; Visentin et al., 2019), whereas phenotypic and genetic correlations between P and lactose content were weak (Costa et al., 2019).

**Other Minerals.** In addition to the abovementioned minerals, 7 other minerals were studied: Cu, Fe, K, Mg, Mn, Se, and Zn. The minerals Cu, Fe, Mn, Se, and

Zn are considered as trace elements, which are present at low concentrations. These minerals are usually associated with milk components such as caseins, whey proteins, citrate, inorganic phosphate, or lactoferrin (Gaucheron, 2013). To the best of our knowledge, no studies have presented results regarding infrared-based prediction of the milk minerals Cu, Fe, Mn, Se, or Zn.

The milk minerals K and Mg, on the other hand, are present at higher concentrations. Minerals do not have specific absorption bands in the infrared spectrum (Huang et al., 2009). For this reason, one could argue that prediction of K or Mg in milk is indirect and depends on other milk components. Potassium is present freely in milk and contributes to the osmolarity of milk (Neville, 1995). Prediction of K could depend on milk components that are related to osmotic equilibrium in milk, such as lactose. The MIR-based prediction models for K in milk have been built, though model performance varied from moderate to good among studies (Soyeurt et al., 2009; Stocco et al., 2016; Visentin et al., 2016; Franzoi et al., 2019).

**Prediction with Infrared-Predicted Milk Components.** For each milk mineral for which the infrared prediction model accuracy >0.60 (Ca, Na, and P), 5 additional prediction models were built based on infrared-predicted fat, protein, and lactose (Table 2). As expected, prediction accuracies of fat-, protein-, or lactose-based prediction models were never higher than infrared prediction models. The RMSEP values were up to a factor 2.6 larger. These results indicate that there is additional information in the infrared spectrum on these components that we currently do not use. However, the results also show that there is considerable overlap between the infrared information used to predict milk protein content and P and Ca. Similarly, there is considerable overlap between the infrared information used to predict milk lactose content and Na.

It could be argued that one should avoid predicting phenotypes with predicted phenotypes. Infrared prediction accuracies for milk fat, protein, and lactose are very high (Luinge et al., 1993). This means that most

variation of these milk components can be obtained from the infrared milk spectra. However, the part of variation from the gold standard phenotype that is absent in the infrared milk spectra will be lost. This loss introduces an error to the predictions based on infrared-predicted milk components.

**Differences Across Studies.** The performance of the prediction models in this study differ from those obtained in some other studies. A reason for this could be the variability of the data used (Davies, 2006). Increased spectral variability can increase prediction model accuracy (Davies, 2006; Soyeurt et al., 2009; Soyeurt et al., 2011; De Marchi et al., 2014). Spectral variability can be increased using multiple breeds to build the prediction model (Zaalberg et al., 2020; Soyeurt et al., 2011), or by using a method to select milk spectra that maximizes spectral variability (Soyeurt et al., 2009; Visentin et al., 2016).

**Data Standardization.** Small differences in infrared-predicted mineral concentrations can be expected due to different times at which milk samples were taken, or because of differences between herds or groups of animals. However, for most minerals, mean infrared-predicted mineral concentrations were very different from the training populations and the study populations (Appendix Table A1). These large differences can be explained by the different spectrometers that were used in this study. Milk samples from the population that were used to build the prediction model were analyzed with a Milkoscan FT2 (FOSS Analytical A/S), whereas the study population's milk samples were analyzed with a Milkoscan FT6000 (FOSS Analytical A/S). Due to this dissimilarity of spectrometers, the predicted mineral concentrations for the genotyped study population were different compared with the gold standard mineral values (Appendix Table A1). These differences are the result of a prediction bias (Bonfatti et al., 2017). The prediction bias will not affect the ranking of the cows, in case it is random. The genetic analysis conducted in the current study, consequently, will not be affected by this prediction bias.

### Genetic Parameter Estimation for Predicted Minerals

**Heritability Estimates.** Table 3 shows an overview of the heritability estimates and proportional variation explained by herd-year-season. The  $h^2$  estimates for infrared-predicted milk mineral concentrations were moderate to high, ranging from 0.13 to 0.36 (Table 3). In other studies on infrared-predicted milk minerals,  $h^2$  of milk minerals were higher, varying from 0.32 to 0.56 (Sanchez et al., 2018; Costa et al., 2019; Visentin et al., 2019). For gold standard milk minerals,  $h^2$  estimates

were between 0.20 and 0.72 (van Hulzen et al., 2009; Buitenhuis et al., 2015).

The infrared-predicted mineral with the highest  $h^2$  was P (0.33–0.36). However, our estimates were lower compared with those observed by other studies, where  $h^2$  varied from 0.53 to 0.56 (Sanchez et al., 2018; Costa et al., 2019; Visentin et al., 2019). Infrared-predicted Na had the lowest heritability (0.13–0.17). This is in line with other studies, who also reported lower  $h^2$  estimates for infrared-predicted Na compared with other minerals (Sanchez et al., 2018; Costa et al., 2019; Visentin et al., 2019).

Genetic and environmental effects on milk production traits change during lactation, especially during early lactation (Vanrobays et al., 2016). Accounting for these changes requires a large data set, which was not available for the current study. Furthermore, the current study only included records of cows in mid lactation. For these 2 reasons, we decided not to correct for changes in genetic and environmental effects.

### Quantitative Trait Loci for Infrared-Predicted Milk Minerals

An overview of the position of the lead SNP group of a QTL, and the percentage of the additive genetic variance that was explained by this lead SNP group is shown in Table 4.

**Breed Differences.** Three QTL were observed for both Danish Holstein and Danish Jersey. This regarded a QTL on BTA 1 and on BTA 14, which were associated with infrared-predicted P in both breeds. The other QTL was located on BTA 17, which was associated with infrared-predicted Ca. However, most of the described QTL were only observed for Danish Jersey. The results of the current study provide us with evidence that differences in genetic architecture of infrared-predicted milk minerals exist between Danish Holstein and Jersey. Both the properties of QTL, and breed differences in QTL properties can have an effect on the accuracy of multibreed genomic prediction (Wientjes et al., 2015). The breed differences observed in our study should be

**Table 3.** Heritability and variance explained by herd-year-season (HYS<sup>2</sup>) for infrared-predicted minerals in Danish Holstein (DH) and Danish Jersey (DJ) cows

Mineral	$h^2$		HYS <sup>2</sup>	
	DH	DJ	DH	DJ
Calcium	0.21	0.19	0.09	0.14
Phosphorus	0.36	0.33	0.06	0.16
Sodium	0.17	0.13	0.13	0.18

taken into consideration, when using infrared-predicted minerals in multibreed genomic prediction.

**Calcium.** Ca content in milk is known to be related to protein content in milk (Gaucheron, 2005, 2013). This is supported by the signals from BTA 6, and the strong ability of infrared-predicted protein content to predict Ca in milk (Table 2). The QTL on BTA 6 was located in the genomic region of the *CSN*-cluster. The *CSN*-cluster codes for the milk caseins.

Studies using the gold standard Ca content found a strong association with the *DGAT1* polymorphism on BTA 14 (Buitenhuis et al., 2015; Bovenhuis et al., 2016). This *DGAT1* signal is not observed in the current study, neither was described by the elaborate GWAS study on infrared-predicted minerals by Sanchez et al. (2019). These results illustrate that a GWAS based on infrared-predicted Ca may not be able to identify QTL that would be identifiable when analyzing gold standard minerals, even when  $R^2 > 0.8$ .

**Sodium.** For infrared-predicted Na in milk, very few QTL were observed. The QTL on BTA 6 was previously associated with infrared-predicted K in milk (Sanchez et al., 2019), milk yield, and protein content in milk (Cohen-Zinder et al., 2005). Neither of these genes have been associated with gold standard Na content (Buitenhuis et al., 2015).

**Phosphorus.** The QTL on BTA 1 harbors the gene *SLC37A1*, which has been associated with milk P concentration (Buitenhuis et al., 2015; Kemper et al., 2016; Sanchez et al., 2019). The gene encodes an antiporter that transfers glucose into the milk by moving a phosphorus molecule in the opposite direction (Pan et al., 2011). This QTL only showed up for P when predicted with infrared milk spectra, but not when predicted with infrared-predicted fat and protein (Figure 1). This QTL has also been identified during a GWAS on

individual wavenumbers (Wang and Bovenhuis, 2018). The results of this study and studies such as Wang and Bovenhuis (2018) show that infrared milk spectra contain additional information that is not obtainable from infrared-predicted milk composition traits alone.

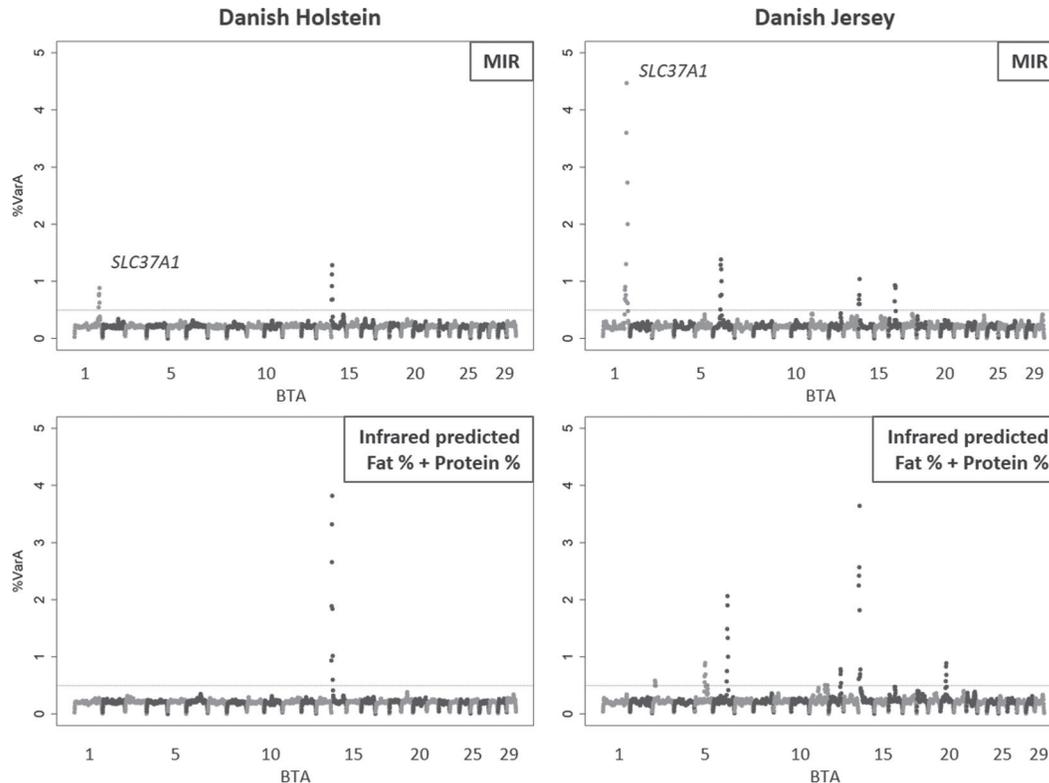
There is a link between P in milk and milk protein content. Organic P is related to the phosphorylation of the caseins and inorganic P (Gaucheron 2005, 2013). Inorganic P is related to the casein micelles as calcium phosphate nanoclusters, or as inorganic phosphorus in the soluble serum phase (Gaucheron 2005). The QTL on BTA 1 and 14 have been associated with infrared-predicted P (Table 4), casein fractions in milk, and cheese-making properties previously (Raven et al., 2014; Sanchez et al., 2018, 2019). In line with this, strong genomic correlation between cheese-making properties and P in milk have been reported (Sanchez et al., 2019).

## CONCLUSIONS

This study has shown that Ca and P in milk can be accurately predicted based on infrared milk spectra, and Na in milk can be moderately predicted. Prediction of milk minerals based on infrared-predicted fat, protein or lactose content is much less reliable than prediction based on the infrared milk spectra, which shows that the full IR spectra contains additional information on milk minerals that is not used at this moment. Infrared-predicted minerals had low to moderate heritabilities. This study also revealed relevant QTL, which provide us with a better understanding of the genetic background of infrared-predicted milk minerals. In some cases, it was possible to detect QTL for gold standard milk minerals that were also associated with infrared-predicted milk minerals, such as

**Table 4.** Lead SNP groups of QTL for Danish Holstein (DH) and Danish Jersey (DJ) for the infrared-predicted milk minerals Ca, Na, and P; only QTL that explained  $\geq 0.5\%$  of the additive genetic variance ( $\sigma_A^2$ ) are presented.

Mineral	Breed	BTA	Upper limit (bp)	Lower limit (bp)	Explained $\sigma_A^2$ (%)	Candidate gene
Ca	DJ	6	81,119,938	88,592,295	0.57	<i>CSN</i> -cluster
	DJ	16	35,969,184	42,625,201	0.71	
	DJ	16	60,428,507	66,313,340	0.56	
	DJ	17	51,024,797	58,285,008	0.66	
	DH	17	52,809,531	57,800,291	0.50	
Na	DJ	6	42,387,759	48,303,392	0.84	<i>ABCG2</i>
	DH	14	1,463,676	5,428,037	0.77	
P	DH	1	143,524,601	147,640,547	0.89	<i>SLC37A1</i>
	DJ	1	141,807,977	146,940,031	7.40	
	DJ	6	42,387,759	48,303,392	1.39	
	DH	14	1,463,676	5,428,037	1.29	
	DJ	14	1,463,676	5,601,692	1.05	
	DJ	16	37,003,329	44,151,335	0.93	



**Figure 1.** Manhattan plots for milk phosphorus predicted with mid-infrared milk spectra (MIR; top) and infrared-predicted fat and protein percentage (bottom) for Danish Holstein (left) and Danish Jersey (right). In the top 2 graphs, the gene *SLC37A1* is indicated.

with *SLC37A1* and infrared-predicted P. In other cases, the infrared-predicted phenotype was not able to detect previously described associations, for example between the *DGAT1* polymorphism and Ca. Finally, this study showed that there are breed differences in genetic architecture of infrared-predicted milk minerals.

### ACKNOWLEDGMENTS

The study was part of the project “Indicators in milk for nitrogen and phosphorus excretion,” financed by the Danish Milk Levy Foundation (Aarhus N, Denmark). The authors have not stated any conflicts of interest.

### REFERENCES

- Bonfatti, V., L. Degano, A. Menegoz, and P. Carnier. 2016. Short communication: Mid-infrared spectroscopy prediction of fine milk composition and technological properties in Italian Simmental. *J. Dairy Sci.* 99:8216–8221. <https://doi.org/10.3168/jds.2016-10953>.
- Bonfatti, V., D. Vicario, A. Lugo, and P. Carnier. 2017. Genetic parameters of measures and population-wide infrared predictions of 92 traits describing the fine composition and technological properties of milk in Italian Simmental cattle. *J. Dairy Sci.* 100:5526–5540. <https://doi.org/10.3168/jds.2016-11667>.
- Bovenhuis, H., I. Jibrila, and J. Dijkstra. 2018. Predicting milk phosphorus content based on genotypic and milk infrared data. Proceedings of the World Congress on Genetics Applied to Livestock Production.
- Bovenhuis, H., M. H. P. W. Visker, N. A. Poulsen, J. Sehested, H. J. F. van Valenberg, J. A. M. van Arendonk, L. B. Larsen, and A. J. Buitenhuis. 2016. Effects of the diacylglycerol o-acyltransferase 1 (*DGAT1*) K232A polymorphism on fatty acid, protein, and mineral composition of dairy cattle milk. *J. Dairy Sci.* 99:3113–3123. <https://doi.org/10.3168/jds.2015-10462>.
- Browning, B. L., and S. R. Browning. 2016. Genotype imputation with millions of reference samples. *Am. J. Hum. Genet.* 98:116–126. <https://doi.org/10.1016/j.ajhg.2015.11.020>.
- Buitenhuis, B., N. A. Poulsen, L. B. Larsen, and J. Sehested. 2015. Estimation of genetic parameters and detection of quantitative trait loci for minerals in Danish Holstein and Danish Jersey milk. *BMC Genet.* 16:52. <https://doi.org/10.1186/s12863-015-0209-9>.
- Cava-Montesinos, P., M. L. Cervera, A. Pastor, and M. de la Guardia. 2005. Room temperature acid sonication ICP-MS multielemental analysis of milk. *Anal. Chim. Acta* 531:111–123. <https://doi.org/10.1016/j.aca.2004.09.093>.
- Cohen-Zinder, M., E. Seroussi, D. M. Larkin, J. J. Loor, A. Everts-van der Wind, J. H. Lee, J. K. Drackley, M. R. Band, A. G. Hernandez, M. Shani, and H. A. Lewin. 2005. Identification of a missense mutation in the bovine *ABCG2* gene with a major effect on the QTL on chromosome 6 affecting milk yield and composition in Holstein cattle. *Genome Res.* 15:936–944. <https://doi.org/10.1101/gr.3806705>.
- Costa, A., G. Visentin, M. De Marchi, M. Cassandro, and M. Penasa. 2019. Genetic relationships of lactose and freezing point with minerals and coagulation traits predicted from milk mid-infrared spectra in Holstein cows. *J. Dairy Sci.* 102:7217–7225. <https://doi.org/10.3168/jds.2018-15378>.
- Davies, A. N. 2006. Back to basics: Calibration statistics. *Spectrosc. Eur.* 18:31–32.

- De Marchi, M., M. Penasa, A. Zidi, and C. L. Manuelian. 2018. Invited review: Use of infrared technologies for the assessment of dairy products—Applications and perspectives. *J. Dairy Sci.* 101:10589–10604. <https://doi.org/10.3168/jds.2018-15202>.
- De Marchi, M., V. Toffanin, M. Cassandro, and M. Penasa. 2014. Invited review: Mid-infrared spectroscopy as phenotyping tool for milk traits. *J. Dairy Sci.* 97:1171–1186. <https://doi.org/10.3168/jds.2013-6799>.
- Deeth, H. C., and M. J. Lewis. 2015. Practical consequences of calcium addition to and removal from milk and milk products. *Int. J. Dairy Technol.* 68:1–10. <https://doi.org/10.1111/1471-0307.12188>.
- Dehareng, F., C. Delfosse, E. Froidment, H. Soyeurt, C. Martin, N. Gengler, A. Vanlierde, and P. Dardenne. 2012. Potential use of milk mid-infrared spectra to predict individual methane emission of dairy cows. *Animal* 6:1694–1701. <https://doi.org/10.1017/S175173112000456>.
- Eskildsen, C. E., M. A. Rasmussen, S. B. Engelsen, L. B. Larsen, N. A. Poulsen, and T. Skov. 2014. Quantification of individual fatty acids in bovine milk by infrared spectroscopy and chemometrics: Understanding predictions of highly collinear reference variables. *J. Dairy Sci.* 97:7940–7951.
- Franzoi, M., G. Niero, M. Penasa, and M. De Marchi. 2019. Development of infrared prediction models for diffusible and micellar minerals in bovine milk. *Animals (Basel)* 9:430. <https://doi.org/10.3390/ani9070430>.
- Gaucheron, F. 2005. The minerals of milk. *Reprod. Nutr. Dev.* 45:473–483. <https://doi.org/10.1051/rnd:2005030>.
- Gaucheron, F. 2013. Chapter 9: Milk minerals, trace elements, and macroelements. Pages 172–185 in *Milk and Dairy Products in Human Nutrition*. Y. W. Park and G. F. W. Haenlein, ed. John Wiley & Sons, Ltd.
- Gebreyesus, G., M. S. Lund, B. Buitenhuis, H. Bovenhuis, N. A. Poulsen, and L. G. Janss. 2017. Modeling heterogeneous (co)variances from adjacent-SNP groups improves genomic prediction for milk protein composition traits. *Genet. Sel. Evol.* 49:89. <https://doi.org/10.1186/s12711-017-0364-8>.
- Haug, A., A. T. Høstmark, and O. M. Harstad. 2007. Bovine milk in human nutrition—A review. *Lipids Health Dis.* 6:25. <https://doi.org/10.1186/1476-511X-6-25>.
- Heuer, C., H. J. Luinge, E. T. G. Lutz, Y. H. Schukken, J. H. van der Maas, H. Wilmsink, and J. P. T. M. Noordhuizen. 2001. Determination of acetone in cow milk by Fourier transform infrared spectroscopy for the detection of subclinical ketosis. *J. Dairy Sci.* 84:575–582.
- Huang, Y., B. A. Rasco, and A. G. Cavinato. 2009. Chapter 13: Fruit juices. Page 361 in *Infrared Spectroscopy for Food Quality Analysis and Control*. D. W. Sun, ed. Academic Press of Elsevier.
- Jensen, H. B., N. A. Poulsen, K. K. Andersen, M. Hammershøj, H. D. Poulsen, and L. B. Larsen. 2012. Distinct composition of bovine milk from Jersey and Holstein-Friesian cows with good, poor, or noncoagulation properties as reflected in protein genetic variants and isoforms. *J. Dairy Sci.* 95:6905–6917. <https://doi.org/10.3168/jds.2012-5675>.
- Keenan, T. W., and S. Patton. 1995. Chapter 3: Determinants of milk volume. Page 91 in *Handbook of Milk Composition*. R. G. Jensen, ed. Academic Press of Elsevier.
- Kemper, K. E., M. D. Littlejohn, T. Lopdell, B. J. Hayes, L. E. Bennett, R. P. Williams, X. Q. Xu, P. M. Visscher, M. J. Carrick, and M. E. Goddard. 2016. Leveraging genetically simple traits to identify small-effect variants for complex phenotypes. *BMC Genomics* 17:858. <https://doi.org/10.1186/s12864-016-3175-3>.
- Kent, J. C., P. G. Arthur, and P. E. Hartmann. 2009. Partitioning of calcium in human milk after freezing and thawing. *Dairy Sci. Technol.* 89:187–191. <https://doi.org/10.1051/dst/2008035>.
- Lucey, J. A., and P. F. Fox. 1993. Importance of calcium and phosphate in cheese manufacture: A review. *J. Dairy Sci.* 76:1714–1724. [https://doi.org/10.3168/jds.S0022-0302\(93\)77504-9](https://doi.org/10.3168/jds.S0022-0302(93)77504-9).
- Luinge, H. J., E. Hop, E. T. G. Lutz, J. A. van Hemert, and E. A. M. de Jong. 1993. Determination of the fat, protein and lactose content of milk using Fourier transform infrared spectrometry. *Anal. Chim. Acta* 284:419–433. [https://doi.org/10.1016/0003-2670\(93\)85328-H](https://doi.org/10.1016/0003-2670(93)85328-H).
- McLaren, W., L. Gil, S. E. Hunt, H. S. Riat, G. R. S. Ritchie, A. Thormann, P. Flicek, and F. Cunningham. 2016. The Ensembl variant effect predictor. *Genome Biol.* 17:122. <https://doi.org/10.1186/s13059-016-0974-4>.
- McParland, S., G. Banos, E. Wall, M. P. Coffey, H. Soyeurt, R. F. Veerkamp, and D. P. Berry. 2011. The use of mid-infrared spectrometry to predict body energy status of Holstein cows. *J. Dairy Sci.* 94:3651–3661. <https://doi.org/10.3168/jds.2010-3965>.
- Neville, M. C. 1995. Chapter 2: The structure of milk. Page 82 in *Handbook of Milk Composition*. R. G. Jensen, ed. Academic Press of Elsevier.
- Pan, C.-J., S.-Y. Chen, H. S. Jun, S. R. Lin, B. C. Mansfield, and J. Y. Chou. 2011. *SLC37A1* and *SLC37A2* are phosphate-linked glucose-6-phosphate antiporters. *PLoS One* 6:e23157. <https://doi.org/10.1371/journal.pone.0023157>.
- Poulsen, N. A., F. Gustavsson, M. Glantz, M. Paulsson, L. B. Larsen, and M. K. Larsen. 2012. The influence of feed and herd on fatty acid composition in 3 dairy breeds (Danish Holstein, Danish Jersey, and Swedish Red). *J. Dairy Sci.* 95:6362–6371. <https://doi.org/10.3168/jds.2012-5820>.
- Raven, L.-A., B. G. Cocks, and B. J. Hayes. 2014. Multibreed genome wide association can improve precision of mapping causative variants underlying milk production in dairy cattle. *BMC Genomics* 15:62. <https://doi.org/10.1186/1471-2164-15-62>.
- Rutten, M. J. M., H. Bovenhuis, J. M. L. Heck, and J. A. M. van Arendonk. 2011. Prediction of Beta-lactoglobulin genotypes based on milk Fourier transform infrared spectra. *J. Dairy Sci.* 94:4183–4188. <https://doi.org/10.3168/jds.2011-4149>.
- Sanchez, M. P., M. El Jabri, S. Minéry, V. Wolf, E. Beuquier, C. Laithier, A. Delacroix-Buchet, M. Brochard, and D. Boichard. 2018. Genetic parameters for cheese-making properties and milk composition predicted from mid-infrared spectra in a large data set of Montbéliarde cows. *J. Dairy Sci.* 101:10048–10061. <https://doi.org/10.3168/jds.2018-14878>.
- Sanchez, M. P., Y. Ramayo-Caldas, V. Wolf, C. Laithier, M. El Jabri, A. Michenet, M. Boussaha, S. Taussat, S. Fritz, A. Delacroix-Buchet, M. Brochard, and D. Boichard. 2019. Sequence-based GWAS, network and pathway analyses reveal genes co-associated with milk cheese-making properties and milk composition in Montbéliarde cows. *Genet. Sel. Evol.* 51:34. <https://doi.org/10.1186/s12711-019-0473-7>.
- Soyeurt, H., D. Bruwier, J.-M. Romnee, N. Gengler, C. Bertozzi, D. Veselko, and P. Dardenne. 2009. potential estimation of major mineral contents in cow milk using mid-infrared spectrometry. *J. Dairy Sci.* 92:2444–2454. <https://doi.org/10.3168/jds.2008-1734>.
- Soyeurt, H., P. Dardenne, F. Dehareng, G. Lognay, D. Veselko, M. Marlier, C. Bertozzi, P. Mayeres, and N. Gengler. 2006. Estimating fatty acid content in cow milk using mid-infrared spectrometry. *J. Dairy Sci.* 89:3690–3695. [https://doi.org/10.3168/jds.S0022-0302\(06\)72409-2](https://doi.org/10.3168/jds.S0022-0302(06)72409-2).
- Soyeurt, H., F. Dehareng, N. Gengler, S. McParland, E. Wall, D. P. Berry, M. Coffey, and P. Dardenne. 2011. Mid-infrared prediction of bovine milk fatty acids across multiple breeds, production systems, and countries. *J. Dairy Sci.* 94:1657–1667. <https://doi.org/10.3168/jds.2010-3408>.
- Stocco, G., C. Cipolat-Gotet, V. Bonfatti, S. Schiavon, G. Bittante, and A. Cecchinato. 2016. *Short communication*: Variations in major mineral contents of Mediterranean buffalo milk and application of Fourier transform infrared spectroscopy for their prediction. *J. Dairy Sci.* 99:8680–8686.
- Toffanin, V., M. De Marchi, N. Lopez-Villalobos, and M. Cassandro. 2015. Effectiveness of mid-infrared spectroscopy for prediction of the contents of calcium and phosphorus, and titratable acidity of milk and their relationship with milk quality and coagulation properties. *Int. Dairy J.* 41:68–73. <https://doi.org/10.1016/j.idairyj.2014.10.002>.
- Van der Drift, S. G. A., R. Jorritsma, J. T. Schonewille, H. M. Knijn, and J. A. Stegeman. 2012. Routine detection of hyperketonemia in

- dairy cows using Fourier transform infrared spectroscopy analysis of beta-hydroxybutyrate and acetone in milk in combination with test-day information. *J. Dairy Sci.* 95:4886–4898.
- van Hulzen, K. J. E., R. C. Sprong, R. van der Meer, and J. A. M. van Arendonk. 2009. Genetic and nongenetic variation in concentration of selenium, calcium, potassium, zinc, magnesium, and phosphorus in milk of Dutch Holstein-Friesian cows. *J. Dairy Sci.* 92:5754–5759. <https://doi.org/10.3168/jds.2009-2406>.
- Vanrobays, M.-L., C. Bastin, J. Vandenplas, H. Hammami, H. Soy-eurt, A. Vanlierde, F. Dehareng, E. Froidmont, and N. Gengler. 2016. Changes throughout lactation in phenotypic and genetic correlations between methane emissions and milk fatty acid contents predicted from milk mid-infrared spectra. *J. Dairy Sci.* 99:7247–7260. <https://doi.org/10.3168/jds.2015-10646>.
- Visentin, G., G. Niero, D. P. Berry, A. Costa, M. Cassandro, M. De Marchi, and M. Penasa. 2019. Genetic (co)variances between milk mineral concentration and chemical composition in lactating Holstein-Friesian dairy cows. *Animal* 13:477–486. <https://doi.org/10.1017/S1751731118001507>.
- Visentin, G., M. Penasa, P. Gottardo, M. Cassandro, and M. De Marchi. 2016. Predictive ability of mid-infrared spectroscopy for major mineral composition and coagulation traits of bovine milk by using the uninformative variable selection algorithm. *J. Dairy Sci.* 99:8137–8145. <https://doi.org/10.3168/jds.2016-11053>.
- Wang, Q., and H. Bovenhuis. 2018. Genome wide association study for milk infrared wavenumbers. *J. Dairy Sci.* 101:2260–2272. <https://doi.org/10.3168/jds.2017-13457>.
- Wientjes, Y. C. J., M. P. L. Calus, M. E. Goddard, and B. J. Hayes. 2015. Impact of QTL properties on accuracy of multi-breed genomic prediction. *Genet. Sel. Evol.* 47:42. <https://doi.org/10.1186/s12711-015-0124-6>.
- Wilmink, J. B. M. 1987. Adjustment of test-day milk, fat and protein yield for age, season and stage of lactation. *Livest. Prod. Sci.* 16:335–348. [https://doi.org/10.1016/0301-6226\(87\)90003-0](https://doi.org/10.1016/0301-6226(87)90003-0).
- Zaalberg, R. M., A. J. Buitenhuis, U. K. Sundekilde, N. A. Poulsen, and H. Bovenhuis. 2020. Genetic analysis of orotic acid predicted with Fourier transform infrared milk spectra. *J. Dairy Sci.* 103:3334–3348. <https://doi.org/10.3168/jds.2018-16057>.
- Zaalberg, R. M., N. Shetty, L. Janss, and A. J. Buitenhuis. 2019. Genetic analysis of Fourier transform infrared milk spectra in Danish Holstein and Danish Jersey. *J. Dairy Sci.* 102:503–510. <https://doi.org/10.3168/jds.2018-14464>.

## ORCID

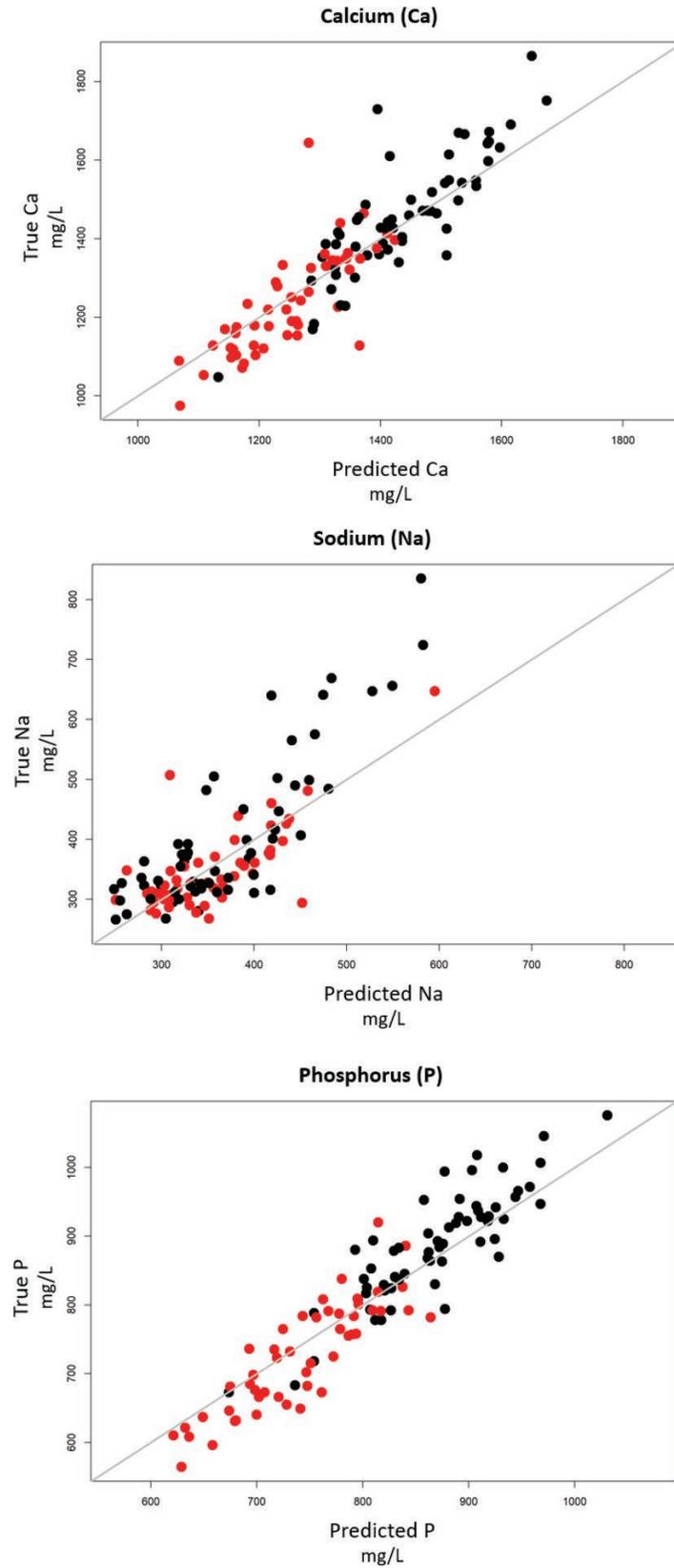
- R. M. Zaalberg  <https://orcid.org/0000-0002-2609-3458>
- N. A. Poulsen  <https://orcid.org/0000-0001-9983-9663>
- H. Bovenhuis  <https://orcid.org/0000-0002-9074-5334>
- J. Sehested  <https://orcid.org/0000-0001-6977-6972>
- L. B. Larsen  <https://orcid.org/0000-0001-9674-0107>
- A. J. Buitenhuis  <https://orcid.org/0000-0002-4953-3081>

## APPENDIX

Table A1. Mean infrared-predicted milk mineral concentrations for the genotyped study population

Mineral	Unit	Danish Holstein		Danish Jersey		Mean infrared-predicted concentration (SD)	
		n	Records	n	Records	Danish Holstein	Danish Jersey
Ca	g/L	1,432	5,666	1,740	7,231	0.53 (0.05)	0.58 (0.06)
Cu	µg/L	1,430	5,686	1,743	7,252	−29.88 (4.38)	−20.23 (4.16)
Fe	mg/L	1,427	5,667	1,736	7,214	0.20 (0.01)	0.22 (0.01)
K	g/L	1,431	5,685	1,739	7,246	2.37 (0.11)	2.60 (0.14)
Mg	mg/L	1,429	5,652	1,744	7,238	0.11 (0.01)	0.12 (0.01)
Mn	µg/L	1,431	5,687	1,748	7,243	−0.98 (2.25)	2.55 (2.38)
Na	g/L	1,433	5,671	1,721	7,201	0.85 (0.04)	0.87 (0.04)
P	g/L	1,433	5,683	1,748	7,280	0.50 (0.04)	0.56 (0.05)
Se	µg/L	1,432	5,678	1,744	7,241	3.73 (0.71)	3.53 (0.74)
Zn	mg/L	1,431	5,693	1,749	7,296	0.94 (0.23)	1.64 (0.25)

Zaalberg et al.: GENETIC ANALYSIS ON INFRARED-PREDICTED MILK MINERALS



**Figure A1.** True versus infrared-predicted milk mineral concentration for the validation population. Red dots represent Danish Holstein cows, and black dots represent Danish Jersey cows.